91602
new 82-10

ontario hydro
research division

COMPLIMENTARY COPY
Please Credit Ontario Hydro Research If
Reproduced in Whole or in Part

To: Mr. A.J. Harris
Manager
Resource Utilization Department
PSOD

ENERGY PLANNING MODEL ENHANCEMENTS

Report No. 88-93-K

L. Wang, Senior Engineer
N. Ramani, Engineer
Reliability & Statistics Section
Operations Research Department

ABSTRACT

One of the tools used in the BESTOP (Bulk Electricity System Tactical Operating Plan) process is the FOMOD program which employs the Monte Carlo simulation technique for electric energy production planning. This report describes two proposed methods of estimating the confidence bounds on the simulation results. It also presents the results of an investigation on possible improvements to the generating unit model used in the simulation process of the FOMOD program.

1. INTRODUCTION

The bulk electricity system (BES) of Ontario Hydro is large and complex. The primary concern in operating the BES is to minimize energy production cost while maintaining reliable customer service. In order to attain economic operations, an elaborate planning process known as BESTOP (Bulk Electricity System Tactical Operating Plan) has been conceived and is described in a CIGRE paper[1]. The models and programs for the implementation of this planning process are currently under development in the Power System Operations Division. The energy planning por-

89-02573/

LEGAL NOTICE

This report was prepared by Ontario Hydro and, while effort has been made to assure the accuracy and applicability of the material contained herein, neither Ontario Hydro nor any person acting on its behalf:

(a) makes any warranty, express or implied, with respect to the use of any information, apparatus, method or process disclosed in this report or that such use may not infringe privately owned rights; or

(b) assumes any liabilities with respect to the use of, or for damages resulting from the use of, any information, apparatus, method or process disclosed in this report.

tion of BESTOP consists of 4 models which are designed to represent hydraulic, nuclear and fossil generation resources and the interface flows of the transmission network[2]. This report describes enhancements proposed for the fossil generation model (FOMOD).

The FOMOD program is used to generate forecasts of electrical energy production from generating units and stations. The operation of the fossil-fired generation system is simulated by the Monte Carlo method for a given time period under specified conditions. The program output contains the average values of the simulated energy production of individual generating units and stations for this period, and these average values are the point estimates of the required energy production forecasts.

The Monte Carlo simulation approach relies on a direct computer imitation of generation system operations, which recognizes the chronology of all simulated events. Each simulated operation of the generation system over a specified study interval represents one possible realization of system operation. Results of different realizations obtained by this technique may vary over a wide range which provides additional insights on possible "good", "poor" and average performances of the generation system.

It is well known that the output of a Monte Carlo simulation process is a random variable. A point estimate based on a small number of simulation runs gives incomplete and often inaccurate results. In this application, it translates into inaccurate energy production forecasts. To provide a measure of the forecast accuracy, the confidence bounds of simulation results need to be estimated. Appendix A of this report describes a method of estimating confidence limits using the statistical approach. Another method using the time series analysis approach is given in Appendix B.

At present, the basic model for generating units in the simulation process of the FOMOD program consists of the up (operating) and down (outage) states. The representation of the derated operation state by equivalent full outage state in the generating unit model represents a crude approximation. Since the Monte Carlo simulation method is flexible in handling multi-state generating unit models with an efficiency almost equal to that in the case of the two-state model, an enhancement of the FOMOD program to include derated operations was considered desirable. However, the forecastperformance indices of generating units in the Consistent Energy Set (CES) are based on the two-state model. A method needed to be developed for estimating the parameters of the three-state model from the two-state model data. Such a method and an algorithm are described in Appendix C of this report.

A review of existing Monte Carlo simulation programs for energy production is given in Appendix D.

2. SUMMARY OF RESULTS

Approximate Confidence Bounds

A method of calculating the approximate confidence bounds on the simulation results has been derived from the analysis of data generated by a simplified simulation program. The method is described in more detail in Appendix A, and also in Reference 3. It was found that the simulated monthly operating times of baseload and two-shift units can be approximated by normal distribu-

tions. It was also found that the ratio of the sample standard deviation to the mean (also known as the coefficient of variation) of simulated monthly operating times can be approximated by the follow empirical formula:

$$\sigma/\mu = 0.07\ F\ \sqrt{r} + 0.045 \qquad (1)$$

where

μ = sample mean of the simulation results,

σ = sample standard deviation,

F = DAFOR of the generating unit,

r = mean repair time of the unit, (hours).

The approximate confidence bounds on the result of a sample of n realizations (or n monthly operations) for a generating unit can be estimated as follows. Let

X = the mean of n simulated monthly operating times,

S = standard deviation of the mean.

Then, from well established theory and Equation (1),

$$S = (X/\sqrt{n})\ (0.07F\ \sqrt{r} + 0.045) \qquad (2)$$

The 90% confidence bounds on the mean value of the simulated monthly operating time are, respectively,

$$B_U = X + t_{\alpha/2,n-1}\ S \qquad (3)$$

$$B_L = X - t_{\alpha/2,n-1}\ S \qquad (4)$$

The value of $t_{\alpha/2,n-1}$ in the above equations can be found from the table for the t-distribution. Note that $\alpha/2 = 0.05$ for calculating the 90% confidence bounds.

Confidence Interval Estimates Using Time Series

An alternative method of obtaining the confidence bounds is by treating the observations making up a simulation run as a time series. The method is described in Appendix B and is summarized below. Let the number of observations per run be m and let b be the number of runs. Let $Y_{i,j}$ be the jth observation in the ith run, ie, $Y_{i,j} = Y_j$ for run i. For each run i, define

$$B_i = \Sigma_j\ j\ Y_{i,j} - Y_{i,m}\ \Sigma_j\ j$$

$$= \Sigma_j\ j\ Y_{i,j} - Y_{i,m}\ m\ (m+1)/2 \qquad (5)$$

Then H, the half width of the confidence interval is given by

$$H = t_c \sqrt{[\, 12 \, \Sigma_i \, B_i^2 \, / \, n^2(m^2+1)]} \qquad (6)$$

where

c = the chosen confidence coeffient,

t_c = value of the t variable (with b degrees of freedom) that is exceeded with a probability c (available from standard statistical tables), and

n = mb.

Clearly, for each run B_i can be calculated recursively, ie, without any need for storing all the values $Y_{i,j}$.

A comparison of the above estimator with the classical estimator showed that the above estimator using the time series usually outperforms the classical estimator in providing more accurate and more robust estimates (see Appendix B). However, as the number of runs increases, (say $\geq$ 10) the difference between the two estimates becomes negligible. The most useful feature of the time series estimator is the fact that it can provide confidence bounds even when the simulation consists of only one run (whereas the classical estimator cannot).

Three-State Model Parameter Estimation

An iterative method and an algorithm have been developed for estimating the needed parameters of the three-state model from the forecast value of DAFOR and other available data. This method is based on the following assumptions:

(1) The mean repair times (or transition rates into higher capacity states) are known or can be derived from recorded data.
(2) The average derating performances of generating units in the forecasting period (e.g. the coming year) will not be different from those in the past.

The algorithm for computing the failure rates of the three-state model is simple and easy to implement. Both the algorithm and the procedure for carrying out the simulation using the three-state model are described in Appendix C.

3. CONCLUDING REMARKS

One of the criticisms of the Monte Carlo simulation method is that the users are often not sure of the precision of the simulation results. The proposed method of estimating confidence bounds on the simulation results provides a much needed measurement of output precision. This method can also be used to determine the number of simulation runs needed to achieve a specified degree of precision, resulting in savings in computing cost.

The extension of the FOMOD model to include the derated operations of generating units should be given high priority. The method proposed in this report for estimating the parameters of the

three-state model of generating units offers a simple approach to the implementation of the extension. The algorithm for estimating the three-state model parameters can be implemented as a subprogram in the FOMOD model.

Finally, the review of several Monte Carlo simulation programs for electric energy production planning provides a perspective on Hydro's effort in this area. The programs developed by individual utilities are designed to satisfy their own needs. It is true in general that each utility has its own unique set of operating environment, practices and constraints. It is not likely that a program can be developed with all the options to satisfy the requirements of many utilities and still be able to maintain a high level of computational efficiency and speed.

Approved:

J.G. Cassan
Manager
Operations Research Department

Submitted:

L. Wang
Senior Engineer - Reliability
Reliability & Statistics Section

N. Ramani
Engineer
Reliability & Statistics Section

LW/NR:cmh

REFERENCES

1. Winter, W.H., "BESTOP, An Operational Planning Process", presented at CIGRE Session 30, Group 37, Paris, September 1984.

2. Lee, D.T., and Szeto, L.M., "A System of Energy Models for Operational Planning", Proceedings of the Canadian Electrical Association Spring Meeting, March, 1984.

3. Wang, L., "Approximate Confidence Bounds on Monte Carlo Simulation Results for Energy Production", IEEE PES 1988 Winter Meeting Paper No. 88 WM 215-6, January 31 - February 5, 1988.

APPENDIX A

APPROXIMATE CONFIDENCE BOUNDS

The Monte Carlo simulation approach relies on a direct computer imitation of generation system operations, which recognizes the chronology of all simulated events. Each simulated operation of the generation system over a specified study interval represents one sample point. Several such realizations constitute a sample. A large sample size is required for the sample mean to be a close approximation of the true mean. The FOMOD program requires substantial computing time to run and, to reduce computing costs, in applications only small sample sizes are taken. Users of this program, therefore, need to estimate the precision of the production forecast provided by the program. Moreover, variations in generating unit outage parameters, such as DAFOR and mean repair times, also affect forecast precision, and this effect, too, needs to be considered. An investigation was carried out to determine the effects of generating unit outage parameters on the simulation results and to estimate confidence bounds of the results. The investigation is described in this Appendix.

METHOD OF INVESTIGATION

A statistical approach is used in this investigation. It requires the accumulation of a large amount of simulation results under various operating scenarios with varied unit outage parameters. The method for estimating confidence bounds is derived from these results through data analysis.

Method of Simulation

Since the FOMOD simulation program requires a significant amount of computing time, it is not suitable for direct use in this study. Instead, a simplified simulation program was written to produce the data needed for the investigation. The simplified program uses a sequential simulation process in which a sequence of operating and repair times of a generating unit is drawn randomly from the specified uptime and downtime distributions. This simulation process is equivalent to one in which a random draw is made at the beginning of each hour of the simulated operation to determine the state of a generating unit during that hour.

The simulated operation of a thermal generating unit is modeled by one of the following three operating modes:

1. Baseload operation: The unit is operated all the time except when it is unavailable due to forced outages.

2. Two-shift operation: The unit is operated between 7:00 AM and 11:00 PM on weekdays only, and is shut down the rest of the time. A minimum operating time constraint can be specified, ie, if the anticipated operating time after start-up following repairs is less than a specified minimum, the unit will not be started for that shift.

3. Reserve stand-by: The unit is operated to replace a failed unit or to cover unexpected load increases. Minimum operating and shutdown time constraints can be specified for such units.

In the FOMOD simulation program, generating units are recommitted each day. A given generating unit may not be in the same operating mode during the entire period of simulation study (normally one month). The results of this study are also applicable to units with mixed modes of operations.

Assumptions

In this study, the following assumptions are made:

1. The simulated monthly energy production of a generating unit is directly proportional to the mean value of its simulated monthly operating time. This assumption is valid for base load units or units in the lower end of the loading order. It is believed to be acceptable for other units as a first order approximation. Hence, approximate confidence bounds on energy production can be calculated from the confidence bounds on the simulated mean monthly operating times.

2. Generating units are represented by a two-state model. The effect of derated operations are assumed to be accounted for by the use of the equivalent forced outage rates (DAFOR) of generating units in the simulation process.

3. Uptimes of generating units are exponentially distributed.

4. The effect of downtime distribution of generating units on simulation results is investigated by assuming two different distributions: exponential and log-normal. The minimum uptime and downtime constraints are observed.

5. The effects of starting failures of generating units on the simulation results are negligible.

RESULTS

In this study, the simulated operation of a generating unit for one month is considered as one sample point. Five hundred (500) sample points were obtained for each scenario in which a generating unit with a given set of outage parameters is simulated to operate in a specified mode (baseload, two-shift, or reserve). Simulation results were obtained and analyzed for a large number of scenarios. The findings are summarized as follows:

1. The differences between the simulation results for the two downtime distributions, exponential and log-normal, are negligible.

2. The distribution of the simulated monthly operating times (or monthly energy productions) of a generating unit varies with the unit DAFOR and mean repair time. In addition:

   (a) Monthly operating times for two-shift units are approximately normally distributed. The resulting distribution for a sample generating unit is shown in Figure A.l.

   (b) Monthly operating times for baseload units with large DAFOR's are also approximately normally distributed. For units with small DAFOR's, the probability of no forced outages in one month of simulated operation is high, and the distribution becomes

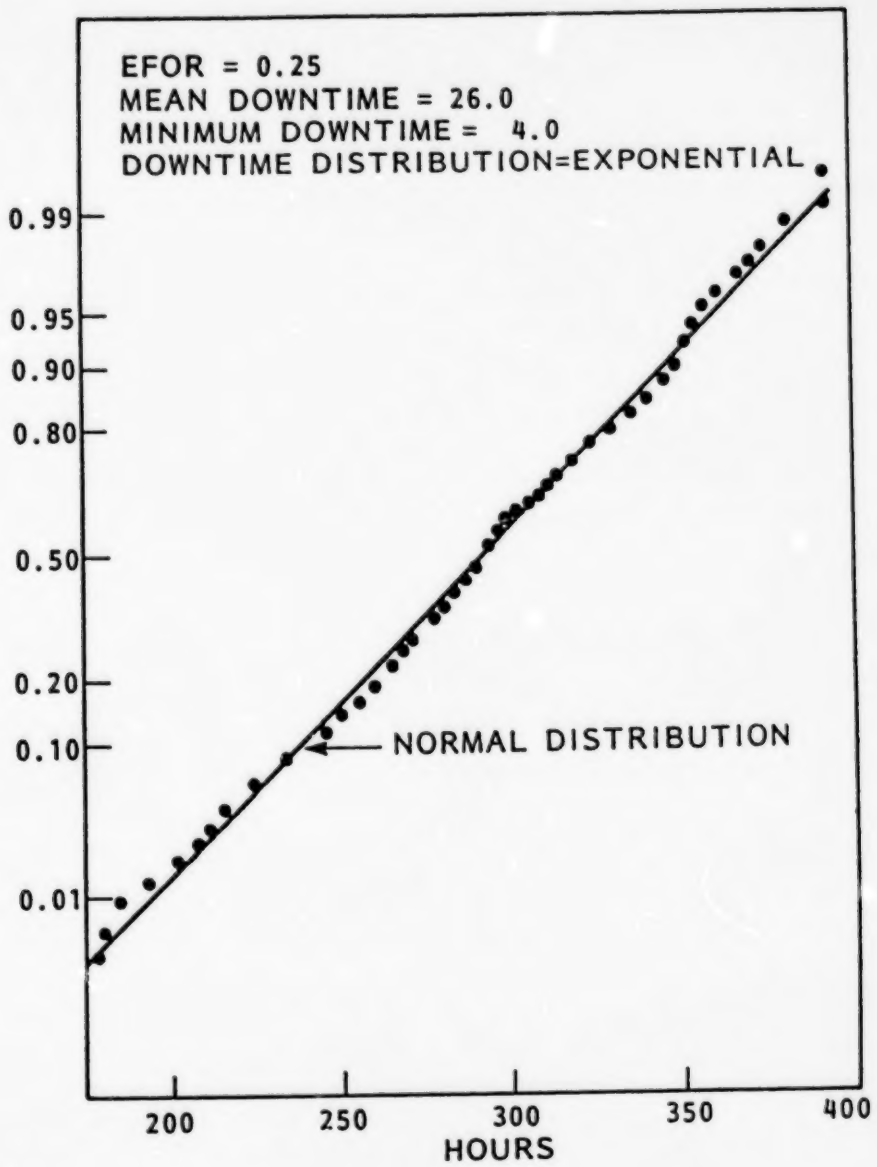

FIGURE A.1

DISTRIBUTION OF SIMULATED MONTHLY OPERATION TIMES - TWO-SHIFT MODE

skewed, limited by the total number of hours in the month (720 or 744) The distribution is shown in Figure A.2.

(c) Monthly operating times for reserve units appear to follow the exponential distribution (see Figure A.3). However, the simulated average monthly operating times of reserve units are very low, and it is felt that their contributions to the overall forecast for the generation system are insignificant. No further investigation was carried out on this mode of operation.

3. It was found that the ratio of the sample standard deviation to the mean (also known as the coefficient of variation) of simulation results can be approximated by the following empirical formula:

$$\sigma/\mu = 0.07\ F\ \sqrt{r} + 0.045 \qquad \text{(A.1)}$$

where

μ = sample mean of the simulation results,

σ = sample standard deviation,

F = DAFOR of the generating unit,

r = mean repair time of the unit, (hours).

This approximation was found to be valid for a unit in the baseload or the two-shift mode of operation over wide ranges of values of DAFOR and mean repair time.

VERIFICATION OF RESULTS

The approximations for the distribution and the empirical formula for the coefficient of variance of simulation results can be verified as follows:

(a) Using Equation (A.1), the standard deviation (σ) is computed from the mean (μ) of the simulated monthly operating times, the DAFOR and the mean repair time of the generating unit.

(b) The 90% confidence bounds of a normal distribution are determined using standard techniques:

$$B_U = \mu + \mu.645\sigma \qquad \text{(A.2)}$$

and

$$B_L = \mu - \mu.645\sigma \qquad \text{(A.3)}$$

(c) Comparisons are made between the 90% confidence intervals obtained by the above computations and the 5 and 95 percentiles of the simulation results.

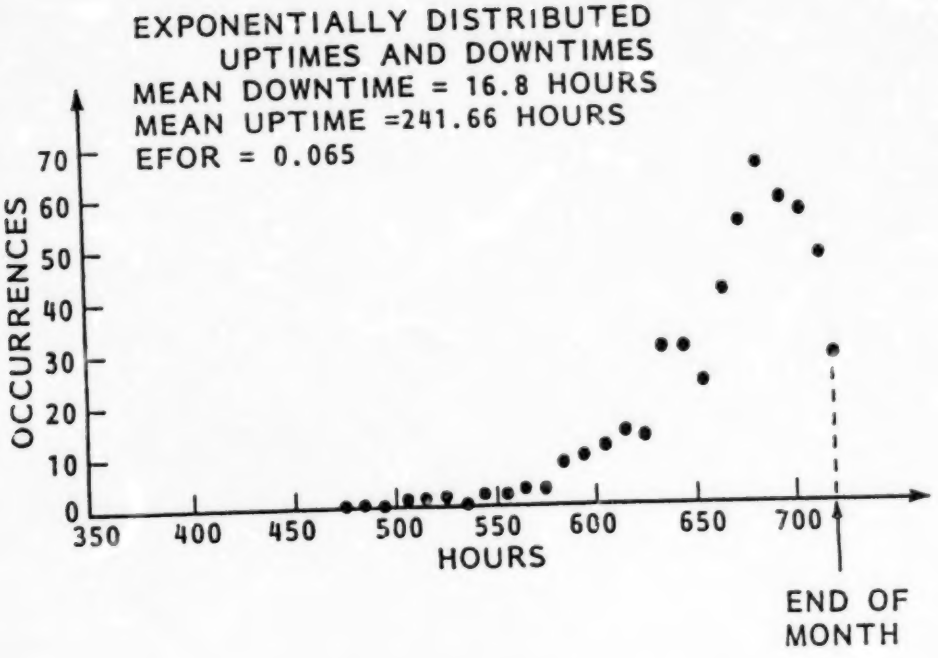

FIGURE A.2

DISTRIBUTION OF SIMULATED MONTHLY
OPERATING TIMES OF BASELOAD UNIT
WITH A SMALL FORCED OUTAGE RATE

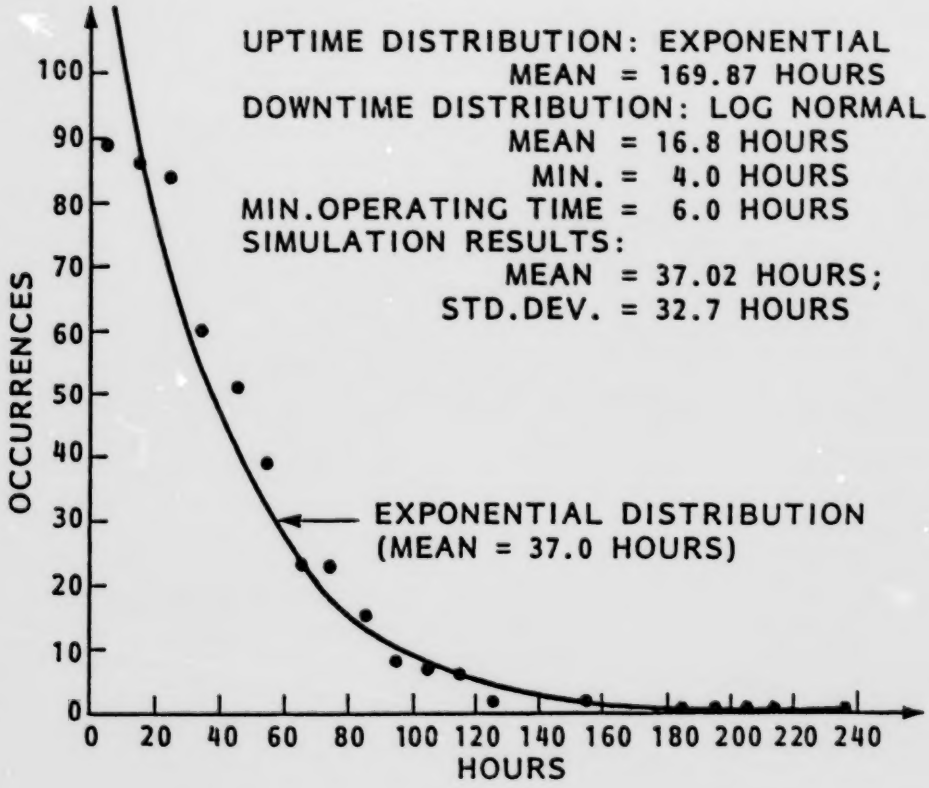

FIGURE A.3

DISTRIBUTION OF SIMULATED MONTHLY
OPERATION TIMES - REVERSE MODE

Results of such comparisons are given in Table A.1 for baseload units and in Table A.2 for two-shift units. The comparisons show that the differences between the calculated confidence intervals and the percentiles of simulation are in the order of 1 or 2% of the mean in most cases, and the largest difference is less than 5%.

Note that for generating units with DAFOR's less than 15%, the upper bound computed by Equation (A.2) may be larger than the maximum possible value of 720 (or 744) hours of operation in a month with no forced outages. In such cases, the maximum possible value is the upper confidence bound.

CALCULATIONS OF CONFIDENCE BOUNDS

An important result of this investigation is that users of Monte Carlo simulation programs for energy production can now estimate the approximate confidence bounds on the result of a sample of n realizations (or n monthly operations) for a generating unit. Let

X = the mean of n simulated monthly operating times,

S = standard deviation of the mean.

Then, from well established theory and Equation (A.1),

$$S = (X/\sqrt{n})\,(0.07F\sqrt{r} + 0.045) \qquad (A.4)$$

The 90% confidence bounds on the mean value of the simulated monthly operating time are, respectively,

$$B_U = X + t_{\alpha/2,n-1}\,S \qquad (A.5)$$

$$B_L = X - t_{\alpha/2,n-1}\,S \qquad (A.6)$$

The value of $t_{\alpha/2,n-1}$ in the above equations can be found from the table for the t-distribution. Note that $\alpha/2 = 0.05$ for calculating the 90% confidence bounds.

Since Equation (A.4) is valid for a unit in baseload or two-shift mode of operation, this equation and the method of computing the 90% confidence bounds can be applied to a generating unit which has been operated in both modes in a given monthly simulation.

For a generating station with m identical units, the mean of station production is equal to the sum of the means of all units in the station, the standard deviation is equal to the square root of the sum of the variances, and the 90% confidence bounds of the station production can be calculated using Equations (A.5) and (A.6).

Equation (A.4) can also be applied to estimate the approximate number of realizations (or sample size) needed to achieve a specified degree of precision. This will be shown in Example 2 below.

TABLE A.1

Comparison of 90% Intervals for Baseload Units

Equivalent forced outage rate	Mean repair time, hrs	Simulated mean operating time, hrs	90% intervals, hrs			
			upper		lower	
			Simulation	Approximation	Simulation	Approximation
0.05	16.8	681.39	720.00	720.00	620.03	614.87
0.09	16.8	654.38	712.24	720.00	578.19	578.14
0.15	16.8	606.56	683.86	694.40	513.21	518.72
0.20	16.8	574.51	658.92	671.27	475.16	477.75
0.25	16.8	540.58	628.57	644.38	444.38	436.78
0.30	16.8	496.04	580.54	602.99	396.54	389.09
0.05	21.4	680.14	720.00	720.00	600.50	611.68
0.09	21.4	656.31	716.00	720.00	569.95	576.25
0.15	21.4	608.41	692.76	702.06	491.97	514.76
0.20	21.4	576.27	672.77	680.32	465.41	472.22
0.25	21.4	537.64	635.87	649.04	419.02	426.24
0.30	21.4	501.22	608.22	618.42	391.44	384.02
0.05	26.0	683.12	720.00	720.00	598.79	612.50
0.09	26.0	656.55	720.00	720.00	564.98	573.25
0.15	26.0	613.88	699.55	713.39	507.53	514.37
0.20	26.0	578.38	677.28	689.11	454.15	467.65
0.25	26.0	535.84	643.83	654.16	413.81	417.52
0.30	26.0	501.96	622.82	627.54	369.35	376.38

TABLE A.2

Comparison of 90% Intervals for Two-shift Units

Equivalent forced outage rate	Mean repair time, hrs	Simulated mean operating time, hrs	90% intervals, hrs upper Simulation	upper Approximation	lower Simulation	lower Approximation
0.05	16.8	344.57	374.63	378.21	309.92	310.93
0.09	16.8	336.46	374.31	375.66	287.79	297.26
0.15	16.8	323.68	361.63	370.56	272.89	276.80
0.20	16.8	315.87	369.98	369.07	250.47	262.67
0.25	16.8	306.37	363.24	365.20	242.42	247.54
0.30	16.8	296.58	361.37	360.53	231.89	232.63
0.05	21.4	341.37	367.64	375.72	303.67	307.01
0.09	21.4	334.75	375.95	375.58	289.42	293.92
0.15	21.4	319.77	370.61	368.99	266.31	270.55
0.20	21.4	310.53	368.84	366.60	251.01	254.46
0.25	21.4	297.53	361.13	359.18	233.63	235.88
0.30	21.4	289.58	351.75	357.29	215.90	221.87
0.05	26.0	340.38	369.21	375.57	300.46	305.19
0.09	26.0	331.55	365.46	373.61	277.52	289.49
0.15	26.0	315.22	362.10	366.32	256.43	264.12
0.20	26.0	308.26	363.54	367.28	244.47	249.24
0.25	26.0	295.07	356.69	360.22	221.73	229.91
0.30	26.0	279.58	350.38	349.52	209.86	209.64

Numerical Examples

Example 1.

Consider a generating station with four identical units each of which is rated at 550 MW and has a DAFOR of 0.264 and a mean repair time of 26.33 hours. The mean monthly operating times for these units obtained from a simulation run of 30 realizations (n=30) are given in Table A.3. The standard deviation of the mean and the 90% confidence bounds on the mean monthly operating time for each unit are computed using Equations (A.5) and (A.6). These values are also given in Table A.3.

TABLE A.3

Confidence Bounds Calculations For Example 1

Unit	Mean monthly operating time, hr	Standard deviation of mean, hr	Approximate 90% confidence bounds Upper, hr	Lower, hr
1	516.89	13.20	539.31	494.47
2	547.22	13.97	570.95	523.48
3	535.17	13.66	558.38	511.96
4	527.45	13.47	550.33	504.47

Assuming that these units are used for base load and are operated at rated capacity all the time, the approximate 90% confidence bounds on the energy production of the entire station can be calculated as follows:

The mean of energy production of the station = $550 \times (516.89 + 547.22 + 535.17 + 527.45)$
$= 1169701.0$ MWh

The standard deviation of the mean = $550 \times (13.2^2 + 13.97^2 + 13.66^2 + 13.47^2)^{1/2}$
$= 25375.73$ MWh

The approximate confidence bounds are calculated, using Equations (A.5) and (A.6), to be 1195077.0 MWh and 1144325.0 MWh, respectively. The coefficient of variation of the mean monthly operating time in this example is 0.026.

Example 2.

This example shows how to estimate the number of realizations required for the generating units in Example 1 so that the coefficient of variation of the mean monthly operating time will not be larger than 0.05 (a milder constraint than that found in example 1). Using Equation (A.4),

$$n = ((0.07 \times 0.264 \times \sqrt{26.33} + 0.045)/0.05)^2 = 7.82$$

Since n must be an integer, select n = 8.

APPENDIX B

CONFIDENCE INTERVAL ESTIMATES USING TIME SERIES DATA

INTRODUCTION

The hourly observations of the state of a generating unit obtained by Monte-Carlo simulation in the FOMOD program form a chronological process, ie a time series. Confidence interval estimation for time series, where the mean value of a parameter is being estimated, has been studied in some detail in the literature [1,2,3]. The methods include the method of replication, batched means and autoregression. A common conclusion appears to be that no one method works best in all cases. Schruben [4] presents a method based on a standardized time series with a minimum of assumptions. This section outlines this method and presents a reformulation to allow the computations to be done recursively. An empirical comparison between this estimator and the classical estimator is then presented.

THE ESTIMATOR

The time series under consideration is assumed to be 'random enough' [4] and strictly stationary, and is given by

$$Y_s = m_0 + e_s, \quad s=1,2,... \tag{B.1}$$

where Y_s is the sth observation, $m_0 = E[Y_s]$ and e_s is a zero mean stochastic process. For the purposes of the current study, it is assumed that the above time series is the output of the simulation, ie, Y_s is the observation at time s and an estimate of m_0 is being sought. Clearly, this estimate is given by the mean of the observations. For a given value of the confidence coefficient c ($0<c<1$), confidence interval estimation is the construction of an interval such that m_0 will be within this interval with a probability c.

Schruben's [4] estimates for the confidence interval are defined as follows. Let the number of observations per run be m and let b be the number of runs. Let $Y_{i,j}$ be the jth observation in the ith run, ie, $Y_{i,j} = Y_j$ for run i. Let $Z_{i,j}$ be the cumulative average of the first j observations on the ith run. That is,

$$Z_{i,j} = (1/j)\ \Sigma_j\ Y_{i,j} \tag{B.2}$$

For each run i, for j = 1,2, ... m, define

$$S_i(j) = Z_{i,m} - Z_{i,j} \tag{B.3}$$

$$T_i(j) = j\ S_i(j) / v\ \sqrt{m} \tag{B.4}$$

and

$$A_i = v\ \sqrt{m}\ \Sigma_j\ T_i(j) \tag{B.5}$$

where v is a scaling constant and is a function of the variance of Y_s. The half-width of the con-

fidence interval for a given value of c is given by

$$H = t_c \sqrt{[12 \Sigma_i A_i^2 / n^2(m^2+1)]} \qquad (B.6)$$

where

t_c = value of the t variable (with b degrees of freedom) that is exceeded with a probability c (available from standard statistical tables)

and

$n = mb$

Now, it can be seen that H does not depend on the value of v. Also, a little reflection will show that a recursive formulation for the above computation can be given by the following. For each run i, define

$$B_i = \Sigma_j \, j \, Y_{i,j} - Y_{i,m} \Sigma_j \, j$$

$$= \Sigma_j \, j \, Y_{i,j} - Y_{i,m} \, m \, (m+1)/2 \qquad (B.7)$$

Then H is given by

$$H = t_c \sqrt{[12 \Sigma_i B_i^2 / n^2(m^2+1)]} \qquad (B.8)$$

Clearly, for each run B_i can be calculated recursively, ie, without any need for storing all the values $Y_{i,j}$.

Note that in many cases, the estimation problem will have to be reformulated slightly in order to satisfy Equation (B.1). For example, if the variable of interest is the total energy supplied by a generating unit in a month, Y_s should be defined as the ratio of the output of the unit at time s to the rated output of the unit. The above method then gives an estimate and confidence bounds for the capacity factor of the unit. These values can then be multiplied by the rated output times the total hours per month to obtain the estimate and confidence bounds for the energy suppiled by the unit in the month.

The algorithm to be used is pretty starightforward. For each run i, compute B_i using Equation (B.7). After the desired number of runs, compute H using Equation (B.8). The confidence bounds are given by $m_0 \pm H$.

SAMPLE COMPARISONS

An empirical comparison of the above estimator with the classical estimator based on three simple examples is presented here.

Three measures of evaluating the performance of the confidence interval estimators are used in the literature. The first is a measure of the accuracy of the estimator and is given by the actual frequency with which the true parameter is contained in the confidence interval. Clearly, the optimal value for this measure is equal to c, ie, 0.9 for the 90% confidence intervals and 0.8 for the 80% intervals. A second measure is the mean half-width of the intervals. Clearly, by increas-

ing the half-width, it is possible to obtain improved coverage at the expense of decreased utility of the bounds. This measure thus gives an indication of the efficiency of the estimator, a smaller half-width indicating a more efficient estimator. A third measure is given by the standard deviation of the half-widths. If the estimates tend to be inconsistent, the estimator is less useful. A high value of the standard deviation thus signifies a less useful estimator.

The first example is a simple two state model and the the second and third examples study a three-state model. In all cases, the parameter being estimated is the mean availability of the unit. Clearly, in the examples, the 'true' value of the variable is known and the above mentioned measures of the performances of the estimators can be applied. In Examples 1 and 3, the mean availability is 0.5 and in Example 2, it is 0.7. In the first example, the state of the unit, and therefore each observation, is either 0 or 1. In examples 2 and 3, there are three possible states: 0, 1, and a derated state between 0 and 1 (assumed to be 0.5 in the examples). In each example, the number of runs is varied from 1 to 20. (Typically, FOMOD simulations consist of 1 or 10 runs).

The results are presented in Tables B.1-B.3. In all cases, the estimator using the standardized time series outperforms the classical estimator in providing more accurate and more robust estimates. However, as the number of runs increases (say $\geq$ 10), the difference between the two estimates becomes negligible. The most useful feature of the time series estimator is the fact that it can provide confidence bounds even when the simulation consists of only one run (whereas the classical estimator cannot).

REFERENCES

1. Law, A.M., and Kelton, W.D., "Confidence Intervals for Steady-State Simulations: I. A Survey of Fixed Sample Size Procedures", Operations Research, 1984, Vol. 32, No.6, pp 1221-1239.

2. Schmeiser, B., "Batch Size Effects in the Analysis of Simulation Output", Operations Research", 1982, Vol.30, No.4, pp 556-568.

3. Sargent, R.G., "An Introduction to Statistical Analysis of Simulation Output", Proc. AGARD Symposium on Modeling and Simulation, Paris, 1979.

4. Schruben, L., "Confidence Interval Estimation Using Standardized Time Series", Operations Research, 1983, Vol. 31, No.6, pp 1090-1108.

TABLE B.1

Performance of 80% Confidence Interval Estimators

Example 1

Number of runs	Fraction within bound	Mean half width	S.D. of width	Fraction within bound	Mean half width	S.D. of width
	Time Series Estimator			Classical Estimator		
1	0.807	0.0268	0.0199	-	-	-
2	0.790	0.0126	0.0066	0.795	0.0185	0.0138
3	0.794	0.0093	0.0039	0.799	0.0103	0.0052
5	0.805	0.0067	0.0022	0.815	0.0070	0.0025
10	0.804	0.0045	0.0010	0.801	0.0046	0.0011
20	0.812	0.0032	0.0005	0.809	0.0032	0.0005

Performance of 90% Confidence Interval Estimators

Example 1

Number of runs	Fraction within bound	Mean half width	S.D. of width	Fraction within bound	Mean half width	S.D. of width
	Time Series Estimator			Classical Estimator		
1	0.902	0.0549	0.0408	-	-	-
2	0.900	0.0196	0.0103	0.907	0.0379	0.0284
3	0.891	0.0134	0.0057	0.904	0.0159	0.0080
5	0.891	0.0091	0.0030	0.907	0.0097	0.0035
10	0.905	0.0059	0.0013	0.908	0.0061	0.0014
20	0.896	0.0041	0.0006	0.908	0.0041	0.0007

TABLE B.2

Performance of 80% Confidence Interval Estimators

Example 2

Number of runs	Fraction within bound	Mean half width	S.D. of width	Fraction within bound	Mean half width	S.D. of width
	Time Series Estimator			Classical Estimator		
1	0.807	0.0416	0.0308	-	-	-
2	0.783	0.0194	0.0103	0.794	0.0285	0.0220
3	0.786	0.0143	0.0061	0.785	0.0160	0.0084
5	0.790	0.0104	0.0033	0.825	0.0108	0.0039
10	0.815	0.0070	0.0016	0.824	0.0071	0.0017
20	0.803	0.0049	0.0008	0.800	0.0049	0.0008

Performance of 90% Confidence Interval Estimators

Example 2

Number of runs	Fraction within bound	Mean half width	S.D. of width	Fraction within bound	Mean half width	S.D. of width
	Time Series Estimator			Classical Estimator		
1	0.904	0.0854	0.0631	-	-	-
2	0.890	0.0301	0.0159	0.884	0.0585	0.0452
3	0.879	0.0205	0.0088	0.892	0.0248	0.0130
5	0.895	0.0141	0.0046	0.907	0.0151	0.0055
10	0.912	0.0092	0.0021	0.905	0.0094	0.0023
20	0.911	0.0064	0.0010	0.908	0.0064	0.0011

TABLE B.3

Performance of 80% Confidence Interval Estimators

Example 3

Number of runs	Fraction within bound	Mean half width	S.D. of width	Fraction within bound	Mean half width	S.D. of width
	Time Series Estimator			Classical Estimator		
1	0.801	0.0363	0.0269	-	-	-
2	0.793	0.0173	0.0090	0.821	0.0261	0.0196
3	0.795	0.0128	0.0053	0.801	0.0144	0.0075
5	0.783	0.0091	0.0029	0.817	0.0097	0.0036
10	0.801	0.0061	0.0013	0.808	0.0064	0.0015
20	0.801	0.0043	0.0007	0.806	0.0044	0.0007

Performance of 90% Confidence Interval Estimators

Example 3

Number of runs	Fraction within bound	Mean half width	S.D. of width	Fraction within bound	Mean half width	S.D. of width
	Time Series Estimator			Classical Estimator		
1	0.910	0.0744	0.0552	-	-	-
2	0.897	0.0268	0.0139	0.895	0.0536	0.0403
3	0.887	0.0184	0.0076	0.888	0.0223	0.0117
5	0.904	0.0125	0.0040	0.908	0.0135	0.0050
10	0.901	0.0081	0.0018	0.908	0.0084	0.0020
20	0.905	0.0056	0.0009	0.897	0.0057	0.0010

APPENDIX C

THREE-STATE MODEL PARAMETER ESTIMATION

The FOMOD simulation program uses a two-state model (operating, failed) to represent the operating cycles of fossil-fired generating units. The reason for using this simple model is due to the fact that it conforms with the official forecast of DAFOR as generating unit performance indices in the Consistent Energy Set (CES). In reality, fossil-fired generating units are often found to be in derated operations. The use of the two-state model represents an approximation and a major source of inaccuracy in the FOMOD program.

Improvement in simulation accuracy can be achieved if a third state representing the derated operations of a generating unit is added to the two-state model. The flexibility of the Monte Carlo simulation method makes it almost equally efficient in applications using the two-state or the three-state model. Therefore, it is desirable that methods be developed to incorporate the three-state model in the FOMOD program. However, it is expected that DAFOR will continue to be the generating unit performance parameter in Ontario Hydro's official forecast. Hence, the success of incorporating the three-state model into FOMOD hinges on whether the parameters required by this model can be derived from the forecast DAFOR value. A method for deriving these parameters from the given DAFOR is described in this Appendix.

THE THREE-STATE MODEL

Only 2 parameters are needed in Monte Carlo simulation using the two-state model. In contrast, the three-state model, shown in Figure C.1, requires 6 parameters (or transition rates). Note that in this figure, the symbol λ is used for transitions from higher to lower capacity states, and the symbol μ is used for transitions from lower to higher capacity states.

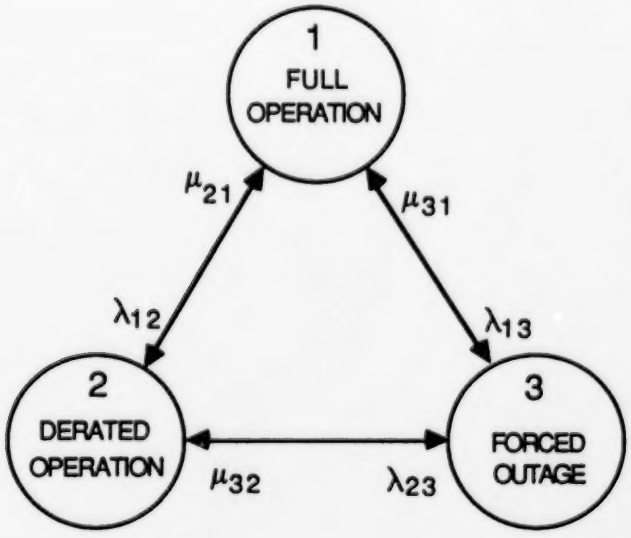

FIGURE C.1 THE THREE-STATE MODEL

ASSUMPTIONS

The proposed method requires the following assumptions:

1. The transition rates into higher capacity states (μ_{21}, μ_{31}, and μ_{32} in Figure C.1) are known. These transition rates are associated with repair efforts and can be derived from recorded data. This assumption is similar to the one for the two-state model regarding the mean repair time.

2. The relative contributions to DAFOR due to forced outages and deratings of a generating unit in the forecast period remain the same as in the past. In other words, in the period of interest the ratio:

$$\frac{\text{equivalent outage time of all deratings}}{\text{forced outage time}} = C_1 \text{ (a constant)}$$

3. The average percentage derating of a generating unit in the forecast period remain the same as in the past. In other words, in the given period of interest the ratio:

$$\frac{\text{equivalent outage time of all deratings}}{\text{derated operation time}} = C_2 \text{ (a constant)}$$

The constants C_1 and C_2 can also be derived from recorded dada.

ALGORITHM FOR COMPUTING FAILURE RATES OF THE THREE-STATE MODEL

The input data for this computing algorithm are the repair rates μ_{21}, μ_{31}, and μ_{32}, the constants C_1 and C_2, and the forecast value of unit DAFOR. Let

P_f = probability of the forced outage state,

P_d = probability of the derated operation state,

P_o = probability of the full operation state.

From the definition of DAFOR, it can be shown that

$$P_f = \text{DAFOR}/(1 + C_1)$$

$$P_d = C_1 P_f / C_2$$

An algorithm was developed using the state probabilities calculated above and the relationships between the state probabilities and the state transition rates for the three-state model of Figure C.1. The steps of this algorithm are as follows:

(1) Compute $P_f = DAFOR/(1 + C_1)$

(2) Compute $P_d = C_1 P_f / C_2$

(3) Compute $P_o = 1 - P_f - P_d$

(4) Compute $\Delta_s = \mu_{21}(\mu_{31} + \mu_{32})$

(5) Compute $\Delta_L = \Delta_s / P_o$

(6) Compute $\lambda_{13} = P_f(\Delta_s + \Delta_L)/2\mu_{21}$

(7) Compute $\lambda_{12} = (\Delta_L P_d - \lambda_{13}\mu_{32})/(\mu_{31} + \mu_{32})$

(8) Compute $\lambda_{23} = (\Delta_L P_f - \lambda_{13}\mu_{21})/(\lambda_{12} + \lambda_{13})$

(9) Compute $\Delta = \lambda_{23}(\lambda_{12} + \mu_{31}) + (\lambda_{12} + \mu_{21})(\mu_{31} + \mu_{32}) + \lambda_{13}(\lambda_{23} + \mu_{21} + \mu_{32})$

(10) Recompute the failure rates by:

$$\lambda_{23} = [\Delta P_o - \mu_{21}(\mu_{31} + \mu_{32})]/\mu_{31}$$

$$\lambda_{13} = (\Delta P_f - \lambda_{12}\lambda_{23})/(\lambda_{23} + \mu_{21})$$

$$\lambda_{12} = (\Delta P_d - \lambda_{13}\mu_{32})/(\mu_{31} + \mu_{32})$$

Note that in the computations using the above formulae, the most up-to-date value of each computed parameter should be used.

(11) Repeat Steps (9) and (10) for as many times as needed to achieve the desired accuracy.

Two sample computations are given in Table C.1. The results show that the computed failure rates converge rapidly to their final values in no more than 4 iterations.

MONTE CARLO SIMULATION USING THE THREE-STATE MODEL

The Monte Carlo simulation process in the FOMOD program calls for hourly random draws for each committed generating unit in the system. The results of the random draws determine the changes of state (from up to down and vise versa) of the generating unit at successive hours in the simulated period of interest.

Simulation using the three-state model can be carried out in a similar manner. In this case, random draws are also made at each hour. The state of a simulated generating unit at the end of an hour depends on its state at the beginning of the hour, the result of the random draw, and the unit transition rates.

The conditional state probabilities are given as follows:

(1) Unit in the up state at t=0.
At t=1,

$$P(up_{t=1}|up_{t=0}) = \exp[-(\lambda_{12} + \lambda_{13})] = p_{11}$$

$$P(derated_{t=1}|up_{t=0}) = (1 - p_{11})\lambda_{12}/(\lambda_{12} + \lambda_{13})$$

$$P(down_{t=1}|up_{t=0}) = (1 - p_{11})\lambda_{13}/(\lambda_{12} + \lambda_{13})$$

(2) Unit in the derated state at t=0.
At t=1,

$$P(derated_{t=1}|derated_{t=0}) = \exp[-(\mu_{21} + \lambda_{23})] = p_{22}$$

$$P(up_{t=1}|derated_{t=0}) = (1 - p_{22})\mu_{21}/(\mu_{21} + \lambda_{23})$$

$$P(down_{t=1}|derated_{t=0}) = (1 - p_{22})\lambda_{23}/(\mu_{21} + \lambda_{23})$$

(3) Unit in the down state at t=0.
At t=1,

$$P(down_{t=1}|down_{t=0}) = \exp[-(\mu_{31} + \mu_{32})] = p_{33}$$

$$P(up_{t=1}|down_{t=0}) = (1 - p_{33})\mu_{31}(\mu_{31} + \mu_{32})$$

$$P(derated_{t=1}|down_{t=0}) = (1 - p_{33})\mu_{32}(\mu_{31} + \mu_{32})$$

From the result of a random draw and the initial condition, one can determine the state of the unit at the end of an hour in the simulation process.

There are two methods of simulating the derated operations. The simpler but less realistic method is to set the level of derating equal to the average level of deratings for this unit on each occasion. The other method employs random sampling from the distribution of deratings for the given generating unit. This distribution must be obtained in advance from recorded data and stored as part of the input data for the unit.

TABLE C.1

Sample Calculations For 3-state Model Parameter Estimation

Sample 1

Input Data	State Probabilities From Input	Estimated Parameters*	State Prob. From Model Parameters
DAFOR=0.2345	P_o=0.6594	λ_{12}=0.02999	P_o=0.6594
C_1=0.1178	P_d=0.1308	λ_{13}=0.00559	P_d=0.1308
C_2=0.189	P_f=0.2098	λ_{23}=0.00323	P_f=0.2098
μ_{21}=0.1586			
μ_{31}=0.013			
μ_{32}=0.0066			

Sample 2

Input Data	State Probabilities From Input	Estimated Parameters*	State Prob. From Model Parameters
DAFOR=0.3015	P_o=0.5761	λ_{12}=0.06045	P_o=0.5761
C_1=0.1045	P_d=0.1509	λ_{13}=0.00544	P_d=0.1509
C_2=0.189	P_f=0.2730	λ_{23}=0.00274	P_f=0.2730
μ_{21}=0.228			
μ_{31}=0.013			
μ_{32}=0.0			

* Results obtained after 4 iterations.

APPENDIX D

A REVIEW OF MONTE CARLO SIMULATION PROGRAMS FOR ENERGY PRODUCTION COSTING

The application of the Monte Carlo simulation method has slowly gained acceptance in the utility industry in North American. In view of its flexibility and ability to mimic generating units and system operations, this is a most suitable method for applications in energy planning and production costing. However, developing a quality simulation program for the application is a major undertaking. Only very few utilities or research organizations have chosen to do so at this time. This Appendix presents the review of existing programs that have come to our attention through published records or personal contact. A tentative comparison of the programs is also attempted. Since most of these programs are proprietary, information on program details are unknown to us.

A feature common to all the Monte Carlo simulation programs reviewed is that random draws are used to determine the chronological operating and outage states of generating units. Once the states of the generating units in the system are determined, economic rules are followed in unit commitment and dispatch.

(1) PRODCOST

Perhaps the first Monte Carlo simulation program for production costing in North America is the PRODCOST program developed by Pennsylvania Power and Light Company. This program is designed to simulate the operation of a generating system consisting of separate regions with inter-tie connections. Within each region, the operations of generating units are simulated by sequences of periods of two hours each. The state of a generating unit in each two-hour period is determined by comparing the result of a random draw with the unit's forced outage rate. Outages of hydraulic units are not simulated, their effects are accounted for by derating the units by the amount equivalent to their forced outage rates.

This program appears to be quite primitive. The two-hour duration for outage states are not realistic and cannot properly consider such factors as ramping rate limitations and minimum downtimes.

(2) BENCHMARK

This is probably the most publicized Monte Carlo simulation program for electrical energy production planning. Its development was sponsored by EPRI and has been under testing for some time. The program simulates the hour-by-hour operation of a generation system in weekly intervals. The emphasis of the simulated operation seems to be on the dispatch process of demand responsive generation, including thermal, conventional hydro and pumped storage hydro units. In the dispatch process, the highest priority is given to reducing capacity shortages and possible capacity surpluses due to steam downturn constraints, followed by economic dispatch.

The main weakness of the program seems to be the simulation of operating and outage states of thermal units. These states are determined by random draws, their durations are in terms of weeks which bear little resemblance to the actual unit operating history.

(3) PACIFIC GAS & ELECTRIC PROGRAM

Recently, Pacific Gas and Electric Company of San Francisco has finished the first phase of the development of a simulation program which does not appear to have the shortcomings of the other two programs. No publication or documentation of the program is available at present. The information obtained through personal contact indicates that the developer of this program placed special emphasis on making the simulation as close as possible to real system operations. Hence, the usual operating constraints of generating units, such as ramping rates and minimum operating times and downtimes are considered, as well as partial outages and outage postponabilities. The simulation is event based and thus the up and down times are not limited to an integral number of hours. It is not clear whether the improved accuracy has a significant impact on the results. Another useful feature of the program is that it has graphical output capabilities. A possible weakness of the program is that it requires a large amount of input data. The derivation of some of the input data from outage records may not be an easy task. The simulation results are sensitive to the input data.

Comparisons of these programs with BESTOP are given in Table D.1.

TABLE D.1

Program Comparison

	BESTOP	PRODCOST	BENCHMARK	PG & E
1. Gen. Unit Model				
A. Thermal				
• Partial outage states	Not yet	No	Yes	Yes
• Random draw of unit state	Hourly	Bi-hourly	Week-by-week	As required by events
• Ramp rate restrictions	Yes	No	Yes	Yes
• Min shutdown time	Yes	No	Yes	Yes
• Start-up restrictions	Yes	No	No	Yes
• Must run units	Yes	Yes	Yes	Yes
• Outage postponability	No	No	No	Yes
B. Hydro				
• Forced outages	To be implemented	No	No	Yes
• Energy limit	To be implemented	No	Yes	Yes
• Storage management	To be implemented	No	Yes	Yes
• Max and Min energy/week	To be implemented	No	Yes	-
• Pumped storage	To be implemented	No	Yes	Yes
2. Unit Commitment and Dispatch				
• Unit commitment interval	Daily	-	Weekly	-
• Commitment order	Average running cost	-	Unit type, full output cost	-
• Operating reserve recognition	Yes	-	Yes	Yes
• Two shifting	Yes	-	No	Yes
• Out of order dispatch	Yes	-	No	Yes
• Dispatch interval	Hourly	-	Hourly	Hourly
• Dispatch order	Incremental cost	Incremental cost	Incremental cost	Incremental cost
• Plant output limit (Transmission bottling)	Yes	Yes	No	Yes